

ADDING VIRTUAL MEASURING STATIONS TO A NETWORK FOR URBAN AIR POLLUTION MAPPING

A. L. Beaulant^{1*}, G. Perron², J. Kleinpeter², C. Weber³, T. Ranchin¹, L. Wald¹

^{1*}Centre Energétique et Procédés, Ecole des Mines de Paris, B.P 207, 06904 Sophia Antipolis Cedex, France, anne-lise.beaulant@ensmp.fr,

²ASPA (Association pour la Surveillance et l'étude de la Pollution atmosphérique en Alsace), 5 rue de Madrid, 67300 Schiltigheim, France

³Laboratoire Image et Ville, UMR 7011, Université Louis Pasteur, 3 rue de l'Argonne, 67000 Strasbourg, France

ABSTRACT

Maps of pollutants concentration are usually generated by means of interpolation and extrapolation methods. The quality of the results depends mainly of the number of permanent or temporary measuring stations. This paper deals with a method for the virtual densification of the network of stations. The method creates “virtual measuring stations”. It aims at improving the quality of interpolation by increasing the number of data on pollutant concentration. The virtual stations are determined by the means of a classification method applied to each pixel of the area under concern. Discriminating elements are pollutants emission classes, land cover types, urban morphological indicators created to this purpose and distance to major roads. A first implementation was made for particulate matter (PM) for the city of Strasbourg (France) using thin-plates spline interpolation method in Arcview 9 GIS. The relative Root Mean Square Error decreases from 49% for five input stations down to 11% for the virtual stations.

Key Words: pollution map, virtual measuring station, morphological indicators, thin plates spline, RMSE

1. INTRODUCTION

Many large cities in Europe have acquired a measuring network operated by *ad hoc* agencies in order to monitor and analyze air quality. Air quality agencies inform local authorities, which rely on these studies to take decisions and to inform the population in order to reduce health impact caused by air pollution. European policies in this domain request accurate space-time knowledge of individual or collective exposure to pollutants. Therefore, it is becoming more and more essential to know or model the spatial distribution of pollutants concentration at any time and any place in the city in order to map pollution at very local scale.

To answer this need, agencies in charge of studying air pollution are using two different types of methods, with possible combination of both types. Either they map pollutants concentration with numerical models, either they interpolate concentrations values issued from permanent or temporary measuring ground network. The combination of the two methods is a methodology, which have been already tested in the frame of the AirProche project, giving encouraging results. Modeling pollution variability at local scale involves to solve very complex phenomena and needs many inputs such as meteorological parameters, emissions parameters, boundary conditions. Several models are used to map pollutants concentrations at large scale, i.e. 100-1000 km (Schmidt et al. 2001, Vautard et al. 2003, Monteiro et al. 2005). Others such as STREET and ADMS model pollution at very local scale, i.e. street and are usually of limited geographical areas (McHug et al. 1997a, b). Models require many input parameters to provide accurate results and these parameters are seldom available. For these reasons, practitioners presently generate maps of concentrations by means of interpolation and extrapolation methods. Among those, the most used are the kriging method (Carletti et

al. 2000) or co-kriging and kriging with trend method using explained co-factors (ex. : emissions, land use) and by the “thin plates spline” method (Ionescu et al. 1996, 2000).

The accuracy of the results of the interpolation depends mainly on the number of known measurements input to the method. Majority of cities have a monitoring network composed by a small number of measuring stations. According to Stalker & Dickerson (1962), a network of conventional ground measurements requires at least one sampling station of PM per surface of 2.5 km² to estimate the actual concentration at $\pm 20\%$. For our study area, the city of Strasbourg, whose surface is 306 km², a network of approximately 122 stations dedicated to PM would be needed instead of the 5 current ones. But, increasing the number of ground measuring stations requires important material means and is extremely expensive. Furthermore, maintenance has a considerable cost.

To overcome this problem, we propose a method for increasing the number of measurements input to the interpolator. It consists in a virtual densification of the network. Our study is based on the work of Ung et al. (2001, 2002), who introduced the concept of “virtual stations”. The first part of this paper presents the method. Then, we apply the method to the city of Strasbourg and discuss the results. Finally, future work is sketched.

2. METHOD

The starting point of the concept of “virtual station” comes from observations made by agencies in charge of air quality monitoring. These agencies (such as ASPA in Strasbourg) observed that there are places in the city, which present properties similar to

those of the measuring stations (Weber et al. 2002). We made the assumption that a place having the same aerodynamic features than an actual measuring station, belonging to the same emission class than this station and with similar surroundings, will have the same behavior regarding pollutants circulation and therefore, the same concentration. Such a place is a virtual station.

The major attributes, or features, of a station are annual emission of pollutants, land cover, distance to major roads and morphological characteristics of the surrounding buildings that affect the aerodynamic flow.

The emission cadaster informs about the quantity of pollutants (in weight unit) emitted per year per squared kilometer. In spite of inherent uncertainties, emissions inventories are often used as major input in atmospheric numerical models for prediction of pollution levels (Choi et al., 2006). The land cover plays an important role in the pollutants dispersion. Changes in land cover (e.g., buildings compared to trees) imply changes in aerodynamic roughness length, which in turn modify air flow by creating turbulences (Tennekes & Lumley 1972). It is a key parameter to assess air pollution at small scale (Briggs et al. 2000) and as input in dispersion model (Hasager and Thykier-Nielsen 2001). The distance to major roads allows us to take traffic pollution into account in the classification. It is often used as exposure indicator to traffic intensity (Hoek et al. 2002a, b).

Morphological indicators describe the shape of the space surrounding the station.

Individual buildings, their height, orientation and their arrangement influence wind flows and thus spatial distribution of the pollutants (Turbelin and Gibert 2000). For example, in a very large avenue, air flows disperse pollutants, which decreases

pollutants concentrations. Inversely, in a very confined space, pollutants accumulate.

We have selected two indicators to characterize the space surrounding the station.

The Miller indicator I_M characterizes the circularity of this space. It varies from 0 for a linear shape to 1 for a perfectly circular shape (eq. 1):

$$I_M = \frac{4\pi S}{P^2} \quad (1)$$

where S is the visibility surface of the area and P its perimeter.

The Gravelius indicator I_G is an indicator of compacity. It compares the perimeter of the visibility surface to a perimeter of a disk having the same area (eq. 2):

$$I_G = \frac{P}{2\sqrt{\pi S}} \quad (2)$$

The visibility surface S characterizes the open space around the station. It is computed by a rays throw technique by taking into account the buildings position around the station. It is expressed as follows (eq. 3):

$$S = \sum_{i=1}^{N_{ray}} S_{angle}(i) \quad (3)$$

where $S_{angle}(i)$, is a sector in the direction i (eq. 4):

$$S_{angle}(i) = \frac{\pi D_{angle}^2(i)}{N_{ray}} \quad (4)$$

with $D_{angle}(i)$, the visibility distance computed in the direction i , and N_{ray} is the number of thrown rays. Here, N_{ray} is set to 360.

Prior to the determination of virtual stations and their localization, the urban area is divided into cells by a regular grid. The cell size defines the spatial resolution. In our application, the cell size is set to 10 m. For each cell, the Miller and the Gravelius indicators, the emission class, the land cover type and the distance to major roads are

computed or already known. Each cell containing a measuring station is compared to all others. Only are kept those cells whose features are similar to those of the station; they are “virtual stations” for this measuring station and they are precisely localized in the city. In other words, a classification is performed : each pixel is allotted to a measuring station if similar, or rejected. The method uses spectral analysis tools, namely hashcoding tools (Albuisson 1995).

3. APPLICATION TO THE CITY OF STRASBOURG

3.1 Study area

The city of Strasbourg is located in Eastern France, separated from Germany by the Rhine River. Geographical coordinates are: 48.33° and 7.38°. The number of inhabitants is around 273,000. ASPA is the local agency in charge of the air quality measuring network in the city of Strasbourg and vicinity. Fig. 1 displays a map of the measuring network in the city of Strasbourg and vicinity. ASPA operates 11 stations; 5 of them measure (PM_{10, 13, 2.5}). Atmospheric pollution in Strasbourg is mostly due to motor vehicles since there is no heavily polluting industries in this area except a refinery located in Reichstett in the North. Wind is blowing mostly from South-West or North-East because it is channeled in the Rhin valley (REKLIP 1995). It can bring significant pollution clouds from the industrial region of Rhur in Germany.

3.2 Data used

The morphological indicators Miller and Gravelius were computed for the whole city of Strasbourg (fig. 2). We developed scripts in Arcview 3.2, Geographical Information System (GIS) software to that aim and run them on the geographical database BD

TOPO® - © IGN of the French Institute of Geography. The database is georeferenced and contains a volumetric 3D description of the city for buildings. Each building is described by a polygon and each polygon has attributes such as height (minimum, maximum, mean), area and perimeter... The computation provides two images of Strasbourg whose pixels contain the values of the Miller and Gravelius indicators. The computing time requested was very large (several days). All data used are stored into a GIS. The computation provides two images of Strasbourg whose pixels contain the values of the Miller and Gravelius indicators (Fig. 2). On these images, the black color stands for the zero value and indicates a building (for a building $S=0$) or a non-circular space for the Miller index or a non-compact space for the Gravelius indicator.

Conversely the white color stands for the maximum value of Miller (1) or Gravelius and indicates a circular space or a very compact space. Table 1 gives the indicators values for each five measuring stations.

The land cover map is the CORINE Land Cover (2000) with an initial resolution of 100 m. It contains 44 different types of land cover. It was oversampled to a grid cell of 10 m using a nearest neighbor technique (Fig.3, left).

The emission cadaster is a map in a vector format; each polygon represents 1 km² area. The map was oversampled to 10 m using a bilinear technique. The quantity of pollutants was recoded into a limited number of classes (9). The number of classes is determined so that real stations belong to different classes as possible. A polluting source has an influence till 100 m over a surrounding area (Hewitt and Jackson 2003). A spatial filter of Gaussian type was applied to take this influence into account. The resulting image is no longer the emission register but more a map of annual background pollution (Fig. 3, right).

Distance to major roads is computed for both virtual and real stations thanks to a Arcview 9 GIS tool. The distance to traffic pollution is no more than 50 m for real stations. As Hewitt & Jackson (2003) indicate that a polluting source at ground level has an influence till 100 m, we used this 100 m limit to discriminate virtual stations. All cells whose distance to roads exceeds 100 m are rejected. On the figure 4, black triangles are cells that will be removed.

4. RESULTS AND DISCUSSION

Table 2 reports on the number of virtual stations found for each of the five stations measuring PM.

The total number of virtual stations is 635. They are scattered all over the city, which is a true advantage for interpolation method. Compared to the initial 5 stations, one may easily understand the benefit of the virtual densification.

However, we are puzzled in this very first application of the method by the large number of virtual stations for the station STG Nord. Conversely, the number of virtual stations is quite low for two others: STG Clemenceau and STG Est. It means that the selected features are sufficiently discriminating in the last case and maybe not in the first one. More examples are necessary to better understand the creation of the virtual stations.

Mapping the PM concentration from this set of virtual stations has been performed using Arcview 9 GIS tools. Two maps of PM₁₀ concentration were obtained by performing the “thin plates spline” method of the Arcview GIS (Fig. 5). The area is approximately 18 x 28 km² with a spatial resolution of 10 m. Buildings and roads of Strasbourg were laid in the background for better readability. Fig. 5 (left) exhibits the

concentration obtained from the five actual measuring stations (black dots). The low number of measuring points yields to a very uniform pollution map, which is not realistic. The spatial variation of the pollution is closer to mathematical functions used in this method than to reality. Figure 5 (right) is a pollution map obtained from the virtual stations for the same date. Compared to the figure on the left, the map based on virtual stations shows a more realistic pollution distribution. Inside the area delimited by measurements, the representation is very realistic: local variations are visible. However, outside this area, the pollution distribution is less realistic: pollution is uniform, concentrations values are identical over the area. This map has not been validated; but a measuring campaign made in June 2003 indicates that virtual stations preliminary spotted by Ung *et al.* (2001, 2002) were actually behaving like the real stations as expected. Nevertheless the results of this campaign must be interpreted with care since very few virtual stations could be validated by this way. In fact, a limited number of pollution sensor was available. The measurements were made by truck mounted sensors, which limited considerably the places to be monitored.

A quantitative assessment was made as follows. When the pollution map is obtained with the measuring stations exclusively, only four stations over five were used for interpolation. The fifth station is the reference and its concentration is estimated by the interpolation method. The estimated concentration is compared to the actual value. By repeating this procedure (five times, each real stations becoming a reference station), we obtain five errors and compute a root mean square error (RMSE). Finally we take the average RMSE. When the pollution map is obtained adding the virtual stations, only four groups of virtual stations (less than 635 virtual stations) plus their associated real stations are used. By the same way than previously, the real station of the missing group

is the reference, its concentration is estimated and compared to the observed concentration. The process is reiterated five times with virtual stations' groups. Five RMSE are computed and finally, the average is taken. A Fractional Bias is also computed. The general expression for the fractional bias (FB) is given by:

$$FB = 2 \frac{(PR - OB)}{(PR + OB)} \quad (5)$$

where OB and PR refer to the averages of the observed (OB) and predicted (PR) values. It indicates if the results are under or over-predicted.

Table 3 reports on the relative RMSE and the Fractional Bias. It shows that using virtual stations for the mapping improves mapping quality: the relative RMSE decreases from around 49% down to 15%. In other words, the improvement of the quality of the interpolation is around 70% when using the “thin plates spline” method. Fairly similar results are obtained with other interpolation methods (ordinary kriging, Inverse Distance Weighted). Those results confirm the benefit of adding virtual stations and are in agreement with those obtained by Wald et al. (1999). Using a similar approach, in a study on the city of Nantes (France), Wald et al. found that the relative RMSE decreased from 70% (only real stations) to 50% (virtual stations), for the case of black smokes. Moreover values of the Fractional Bias show that the prediction of the concentrations is better when using virtual stations for the interpolation. In this case concentrations are underpredicted with a very small bias whereas they are overpredicted with a larger bias when the mapping is done using exclusively the real stations. However, this method has a limitation. Each real station has a different number of virtual station. The higher the number is, the more it will influence the mapping. Indeed,

many places will have the same concentrations. The number of virtual stations acts as a weight for the interpolation, which tends to smooth very local structures.

3. CONCLUSION

In this paper, we present a method to make denser the network measuring pollution and thus to increase the concentration measurements in order to improve the pollution mapping by interpolation. Making denser consists in virtually adding new measuring stations. Virtual stations are places of a city exhibiting features similar to those of measuring stations. Features are linked to pollution sources (emission register, traffic) and aerodynamical properties of the considered place (morphological indicators, land cover). The approach is not new by itself; this paper offers a strong basis for selecting discriminating features with strong relevance to pollutants circulation. Previous studies by Wald et al. (1999) or Ung et al. (2002) exploited series of satellite data as discriminating features and proved to be sensitive to selected images and to the characteristics of spaceborne sensor.

The first implementation of the method was done for particulates (PM) and for the case study of the city of Strasbourg. From the 5 PM measuring stations initially, we obtained 635 virtual stations homogeneously spread over the area. The benefit of virtual stations to the interpolation is proved in a qualitative way by the pollution map and in a quantitative way by the reduction of the RMSE. Though previous studies support these findings, more cases should be analyzed. Efforts are underway to better analyze the created virtual stations and to validate them. The latter point is not obvious and requests expensive fields campaigns. Moreover it takes several months to carry out a campaign.

Nevertheless this work offers an innovative and strong basis for the development of methods to densify the air quality monitoring network. Technically, the method is easily feasible by practitioners since it involves well-known tools and it uses data they already have. It offers promises to map pollution with high accuracy and thus to represent the pollution distribution variability. This method also appears as complementary to those already existing as virtual stations can constrain air pollution model.

ACKNOWLEDGEMENTS

The authors would like to thank Giath Do'Un - Laboratoire Génie Urbain, Environnement et Habitat (LGUEH), Université Marne la Vallée, Pôle Ville - for computing morphological indicators.

REFERENCES

- Albuisson M., 1995. Optimisation des traitements des images multi-sources. In Comptes-Rendu du 1er colloque international "l'imagerie scientifique et le traitement d'image", Cannes, 4-6 avril 1995, 8-17, Association Aéronautique et Astronomique de France.
- Briggs D., de Hoogh C., Gulliver J., Wills J., Elliott P., Kingham S., Smallbone K., 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments, *The Science of The Total Environment*, Volume 253, Issues 1-3, 15 May 2000, Pages 151-167.
- Carletti R., Picci M., Romano D., 2000. Kriging and bilinear methods for estimating spatial pattern of atmospheric pollutants. *Environmental Monitoring and Assessment*, Vol. 63, 341-359.

Choi YJ, Hyde P, Fernando HJS. Modeling of episodic particulate matter events using a 3-D air quality model with fine grid: applications to a pair of cities in the US/Mexico border. *Atmos Environ* 2006;40(27):5181–201.

Hasager C. B., Thykier-Nielsen S., 2001. IRS-1C LISS III land cover maps at different spatial resolutions used in real-time accidental air pollution deposition modelling, *Remote Sensing of Environment*, Volume 76, Issue 3, June 2001, 326-336.

Hewitt C.N., Jackson A., 2003. *Handbook of Atmospheric Science. Principles and Applications*. Blackwell Publishing, Oxford, UK, 149.

Hoek G., Meliefste K., Cyrys J., Lewné M., Bellander T., Brauer M., Fischer P., Gehring U., Heinrich J., van Vliet P., Brunekreef B., 2002. Spatial variability of fine particle concentrations in three European areas, *Atmospheric Environment*, Volume 36, Issue 25, September 2002, 4077-4088.

Hoek G., Brunekreef B., Goldbohm S., Fisher P., van den Brandt P., 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *The Lancet*, Vol. 360, October 19, 2002, 1203-1209.

Ionescu A., Mayer E., Colda I., 1996. Méthodes mathématiques pour estimer le champ de concentration d'un polluant gazeux à partir des valeurs mesurées aux points dispersés. *Pollution Atmosphérique*, janvier-mars 1996, pp. 78-89.

Ionescu A., Candau Y., Mayer E., Colda I., 2000. Analytical determination and classification of pollutant concentration fields using air pollution monitoring network data. Methodology and application in the Paris area, during episodes with peak nitrogen dioxide levels. *Journal of Environmental Modelling & Software* 15, 565-573.

Monteiro A., Vautard R., Borrego C., Miranda A.I., 2005. Long-term simulations of photo oxidant pollution over Portugal using the CHIMERE model. *Atmospheric Environment*, Vol. 39, Issue 17, June 2005, 3089-3101.

McHugh C. A., Carruthers D. J., Edmunds H. A., 1997. ADMS-Urban: an air quality management system for traffic, domestic and industrial pollution. *International Journal of Environment and Pollution*, Vol. 8, Nos.3-6, 666-675.

McHugh C. A., Carruthers D. J., Edmunds H. A., 1997. ADMS and ADMS-Urban. *International Journal of Environment and Pollution*, Vol. 8, Nos. 306, 437-440.

REKLIP Regio-Klima-Projekt, 1995. Communauté trinationale REKLIP sous la responsabilité scientifique de F. FIEDLER – Atlas climatique du Fossé Rhénan méridional. Vdf Zurich, IFG Offenbach, Ed. Coprur, Strasbourg, France, 240 p.

Schmidt H., Derognat C., Vautard R., Beekmann M., 2001. A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. *Atmospheric Environment*, Vol. 35, 2449–2461.

Stalker W.W., Dickerson R.C., 1962. Sampling station and time requirements for urban air pollution surveys, II, suspended particulate matter and soiling index. *Journal of Air Pollution Control Association*, vol. 12, n°3, 111-128.

Tennekes H., Lumley J.L., 1972. *A First Course in Turbulence*. The Massachusetts Institute of Technology Press, Cambridge, MA, USA, .

Turbelin G., 2000. Modélisation de la turbulence atmosphérique en vue de l'étude du chargement aérodynamique des structures soumises aux effets du vent. Thèse de doctorat Sciences Mécaniques. Université d'Evry Val d'Essonne, France, 183 p.

Ung A., Weber C., Perron G., Hirsch J., Kleinpeter J., Wald L., Ranchin T., 2001. Air pollution mapping over a city - virtual stations and morphological indicators.

Proceedings of 10th International Symposium "Transport and Air Pollution", Boulder, Colorado, USA.

Ung A., Wald, L., Ranchin T., Weber C., Hirsch J., Perron G., Kleinpeter J., 2002.

Satellite data for the air pollution mapping over a city – The use of virtual station. In: Gérard Begni (Ed.), Proceedings of the 21th EARSeL Symposium, Paris, France, A. A. Balkema, Lisse, Abingdon, Exton (PA), Tokyo, 147-151.

Vautard R., Martin D., Beekmann M., Drobinski P., Friedrich R., Jaubertie A., Kley D., Lattuati M., Moral P., Neininger B., Theloke J., 2003. Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model. Journal of Geophysical Research, Vol. 108, 8564.

Weber C., Hirsch J., Puissant A., Durrenberger M., Cerciat M., Schnell L., Bronner A.C., 2002. Morphologie urbaine et répartition spatiale des polluants : de la rue à l'agglomération. Application à la Communauté Urbaine de Strasbourg (CUS). Rapport de recherche, programme ACI "Ville" du MENRT, Strasbourg , 153 p.

Wald L, Basly L, Baleynaud JM. Satellite data for the air pollution mapping. In: Nieuwenhuis, Vanghan, Molenaar, editors. Proceedings of the 18th EARSeL Symposium on operational remote sensing for sustainable development", Enschede, Netherlands, 11–14 May 1998; 1999. ISBN 90-5809-029-9, 133–139.

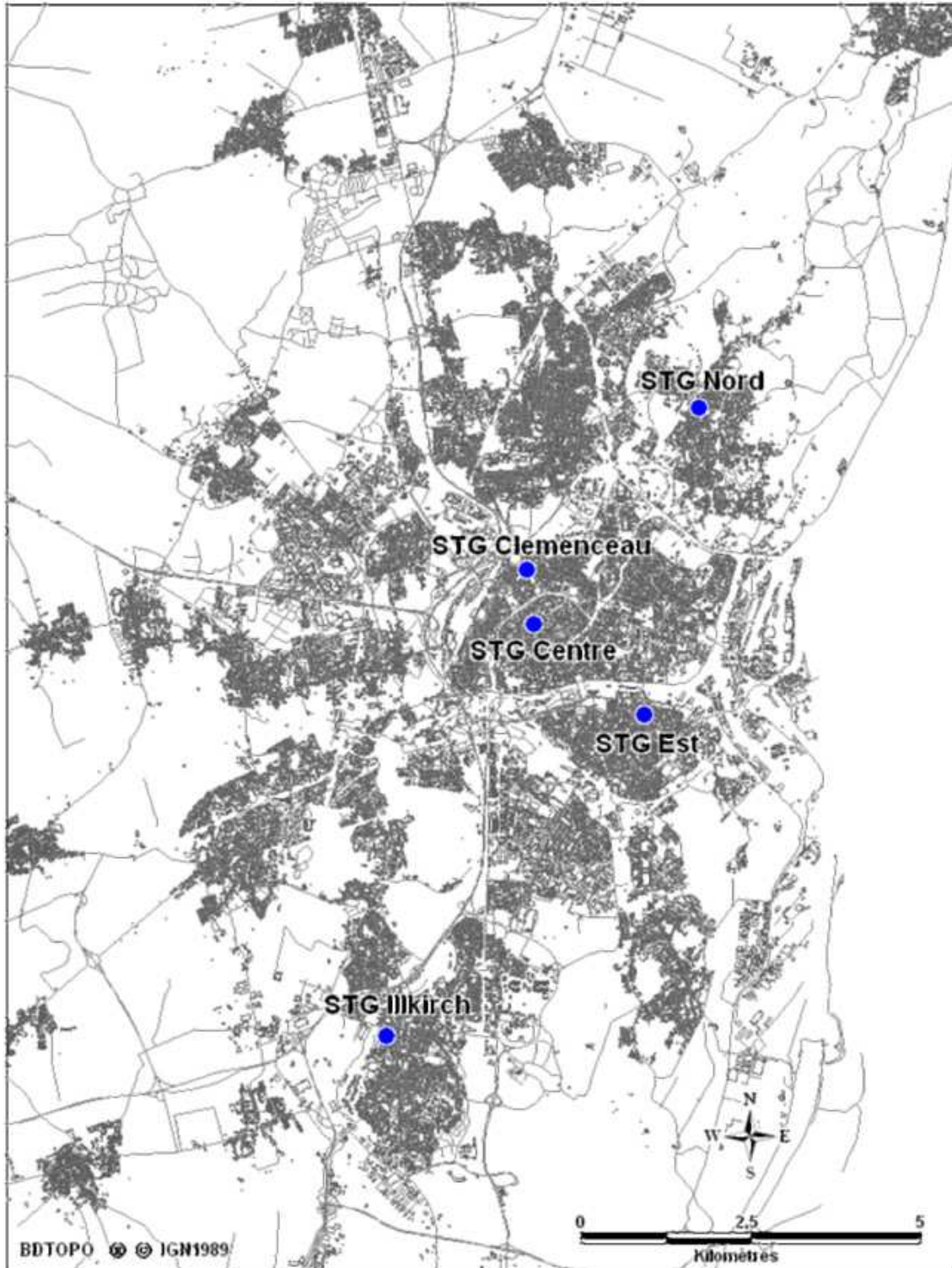


Fig. 1: Stations of the network measuring air quality in the city of Strasbourg. Squares denote stations measuring PM. The width of the map is approximately 20 km.

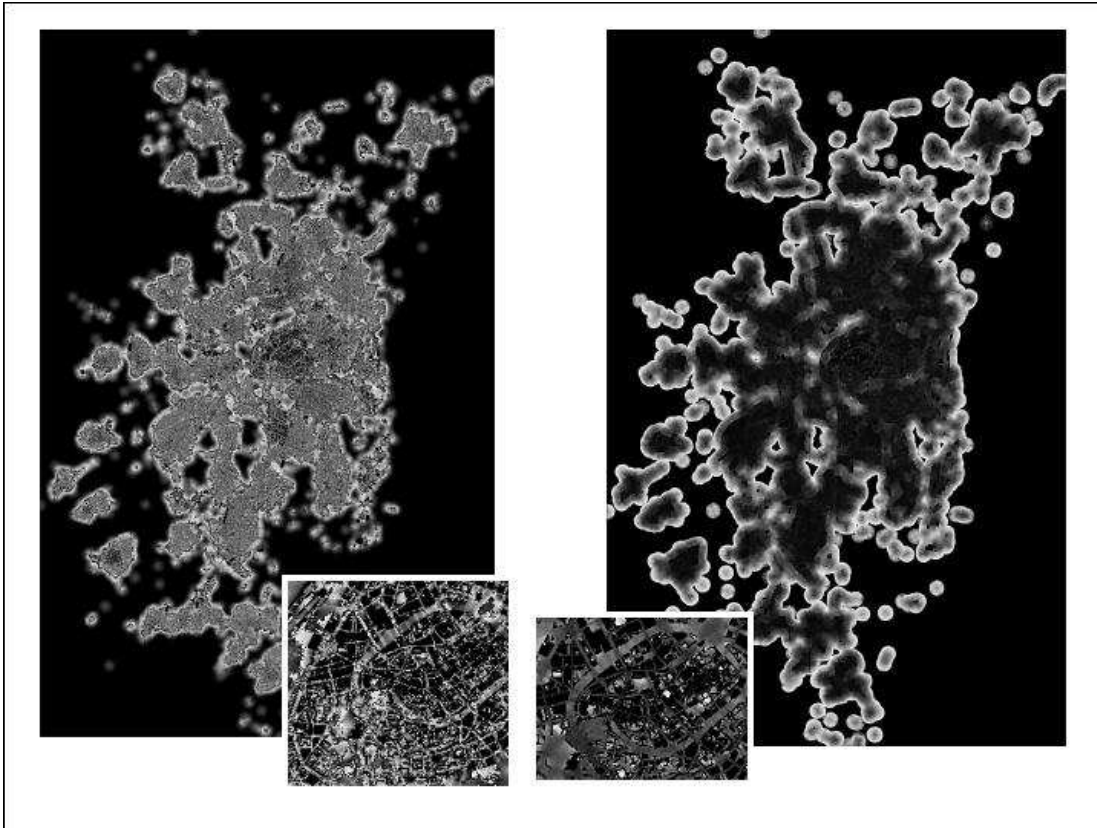


Fig. 2: Images of the Gravelius (left) and Miller (right) indicators for the city of Strasbourg. Images at bottom display zooms of the city center for each indicator. The brighter the gray is, the more compact or circular is the space. The image in the middle is an overview of Strasbourg city (studied area) with buildings in black and roads in black lines.

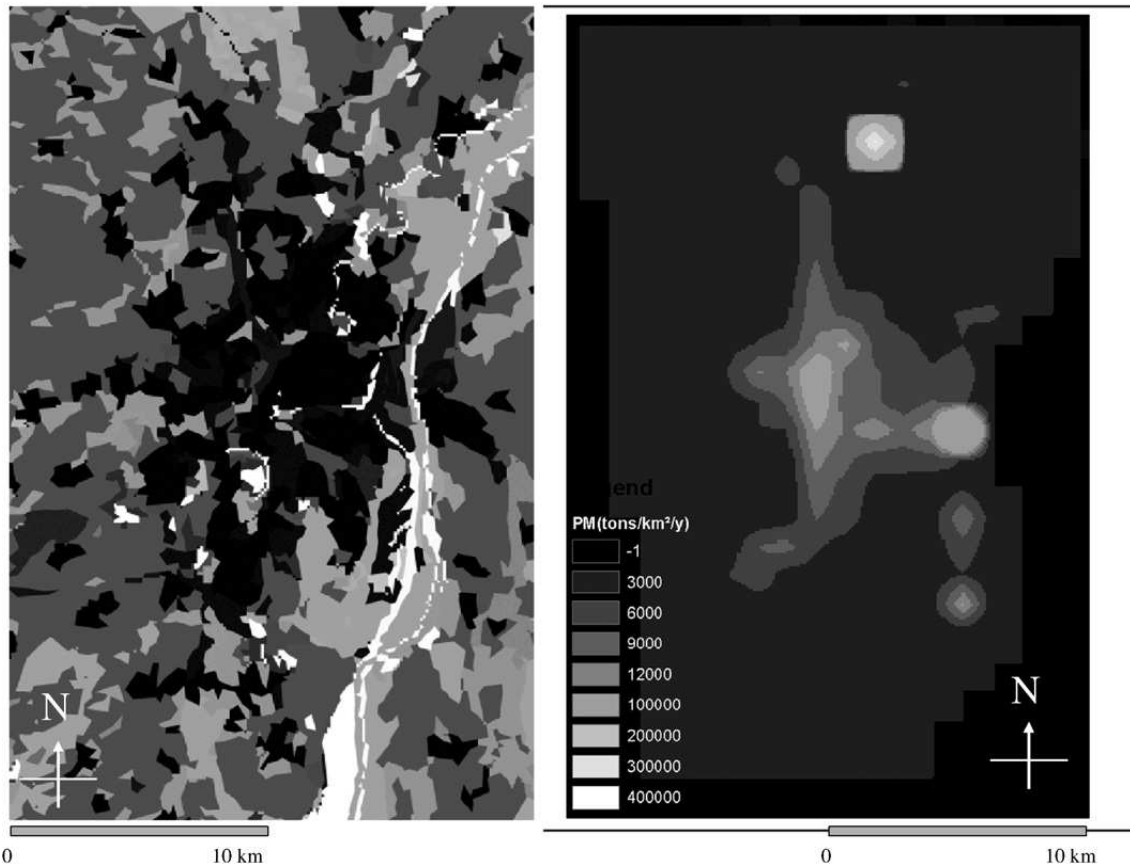


Fig. 3: Images of the land cover map (left) and the emission register for PM for 2004 (right). On the land cover map, the 44 different types of land cover are in different gray levels. For example, urbanized lands are in black.

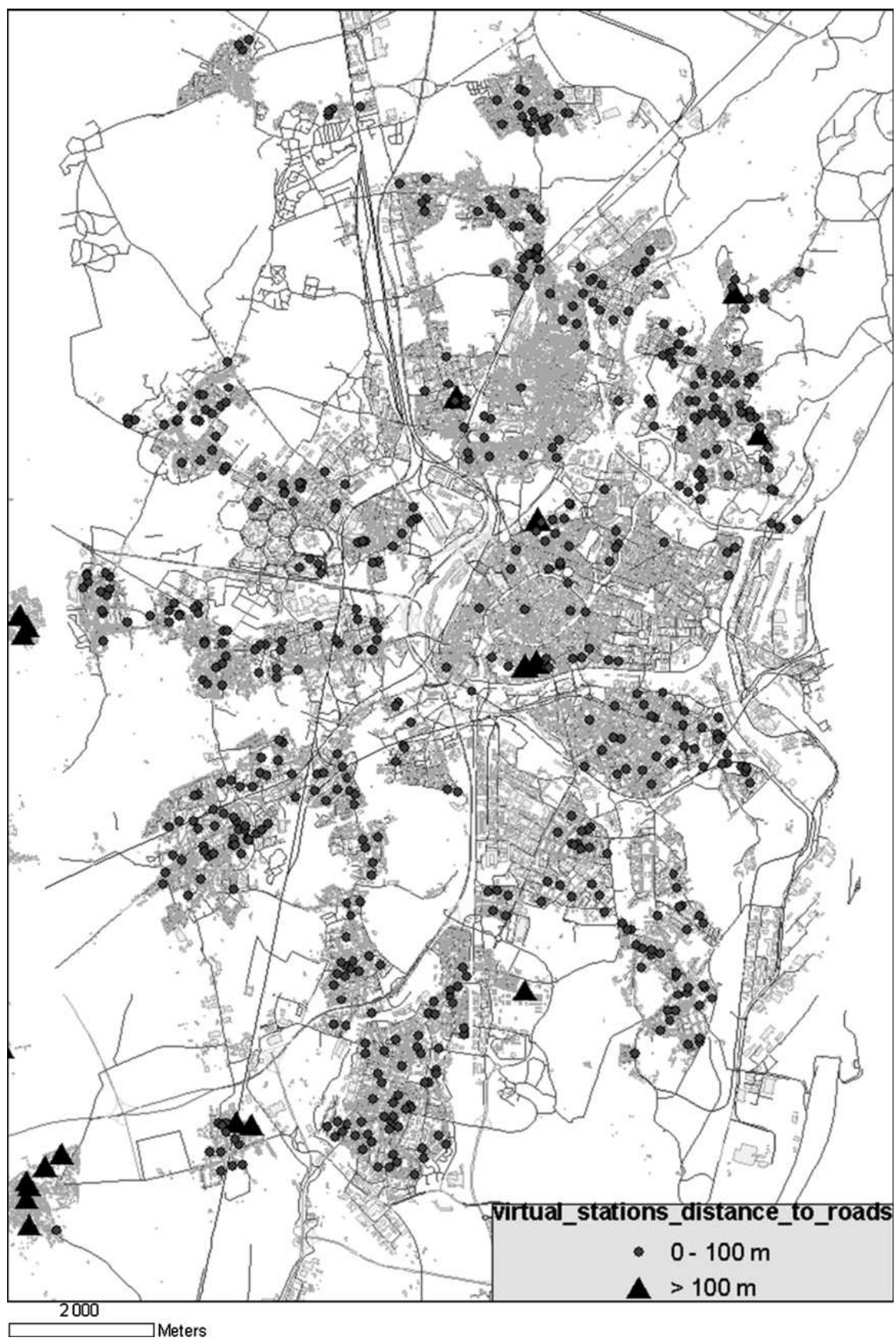


Fig 4 : Virtual stations in Strasbourg. Triangles will be rejected.

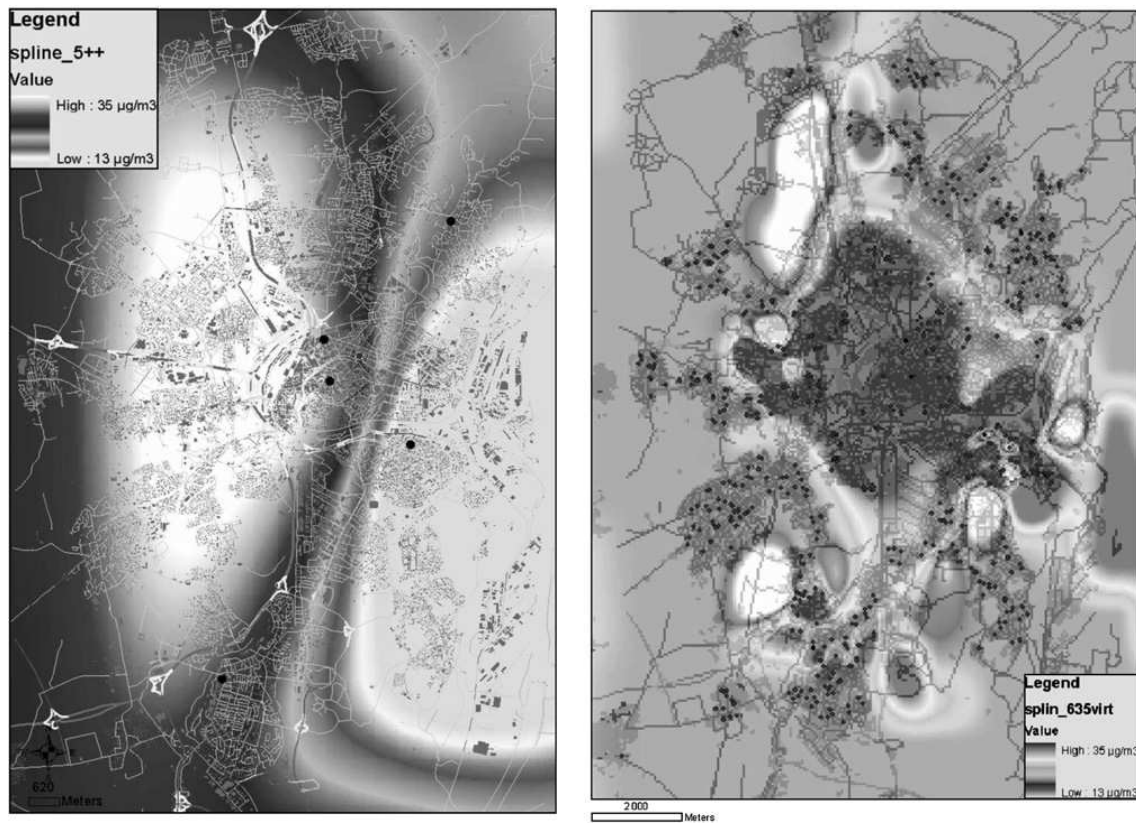


Fig. 5: PM₁₀ concentration maps over Strasbourg obtained by interpolation of real measuring stations only (left, black dots) and of both real and virtual measuring stations (right).

Table 1: Values of Miller and Gravelius for each real measuring station

Real station name	Miller	Gravelius
STG Illkirch	0.07	3.68
STG Centre	0.16	2.54
STG Nord	0.06	4.26
STG Clemenceau	0.07	3.73
STG Est	0.05	4.39

Table 2: Number of virtual stations for each real measuring station

Real station name	Number of virtual stations
STG Illkirch	71
STG Centre	104
STG Nord	450
STG Clemenceau	6
STG Est	4

Table 3: Relative RMSE values and Fractional Bias values for an interpolation with and without virtual measuring stations

“Thin plates spline” method	5 real measuring stations	Real + virtual measuring stations
Relative RMSE	49%	15%
Fractional bias	0.11	-0.07